

Simulation #1 ($N > p$): Examining Variable selection performance of Penalized Case-base

Simulation Settings

We first look at an example of the $N > p$ setting, with N set to 250 and p set to 50. We simulate the covariates X from a multivariate normal distribution ($MVN(0, 1)$) with a correlation matrix of $\rho^{|i-j|}$ where ρ was set to 0.6 for moderately high correlations between variables closer in indices. We consider a sparse setting: For cause 1, The cardinality of the set of betas set to zero was 37, with the non-zero betas taking values of $\{-1, 1\}$. For cause 2, 36 betas were set to zero, with the non-zero betas taking on values of $\{-1, 1\}$.

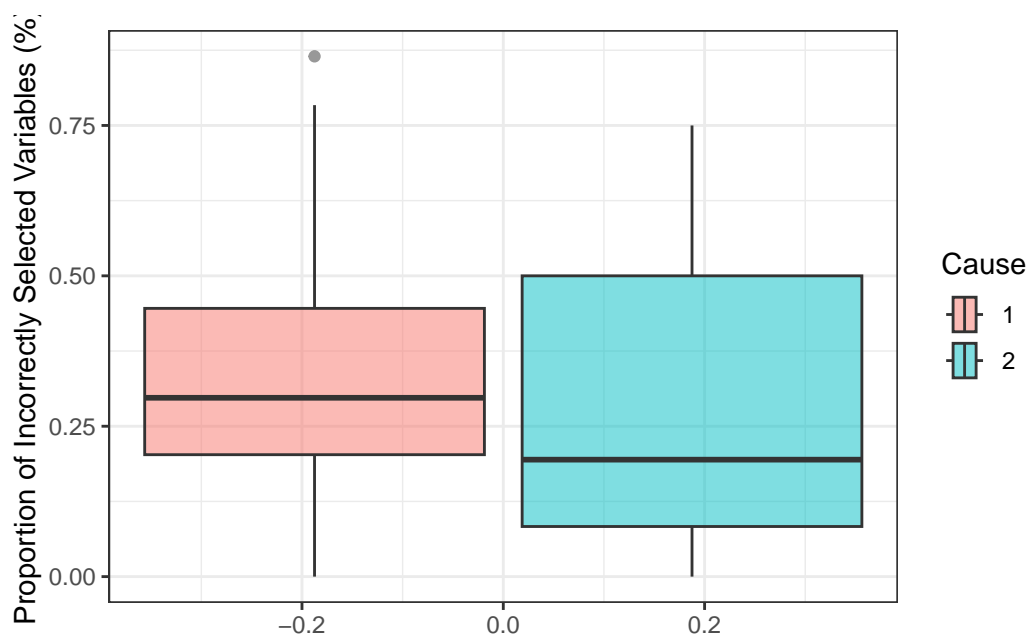
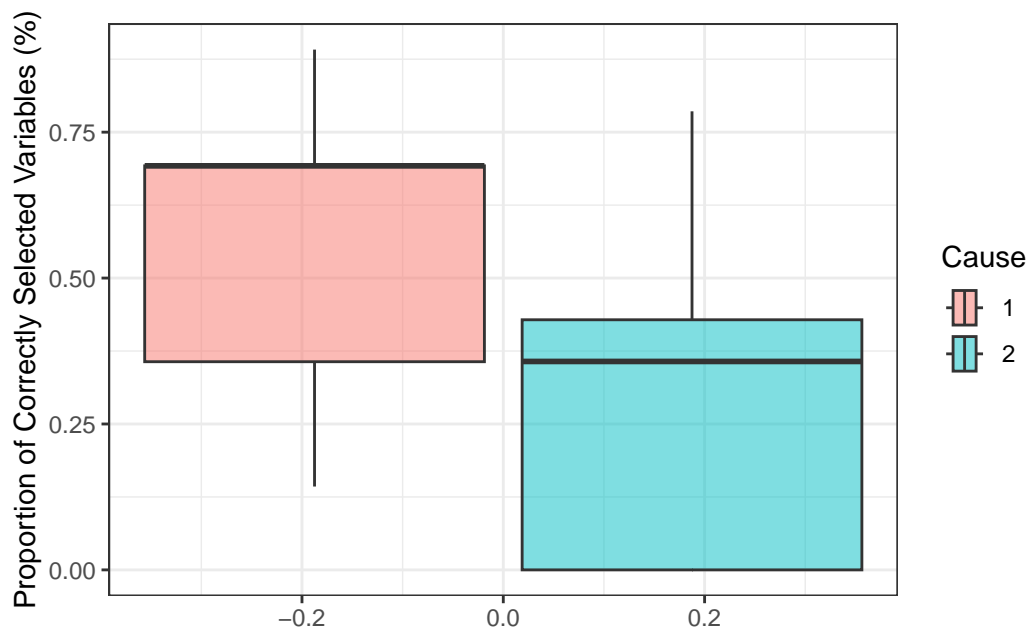
```
beta1 <- c(rep(-1, 4), rep(0, 5), rep(1, 4), rep(0, 16), rep(1, 5), rep(0, 16))
beta2 <- c(rep(0, 4), rep(-1, 5), rep(1, 4), rep(0, 16), rep(1, 5), rep(0, 16))
```

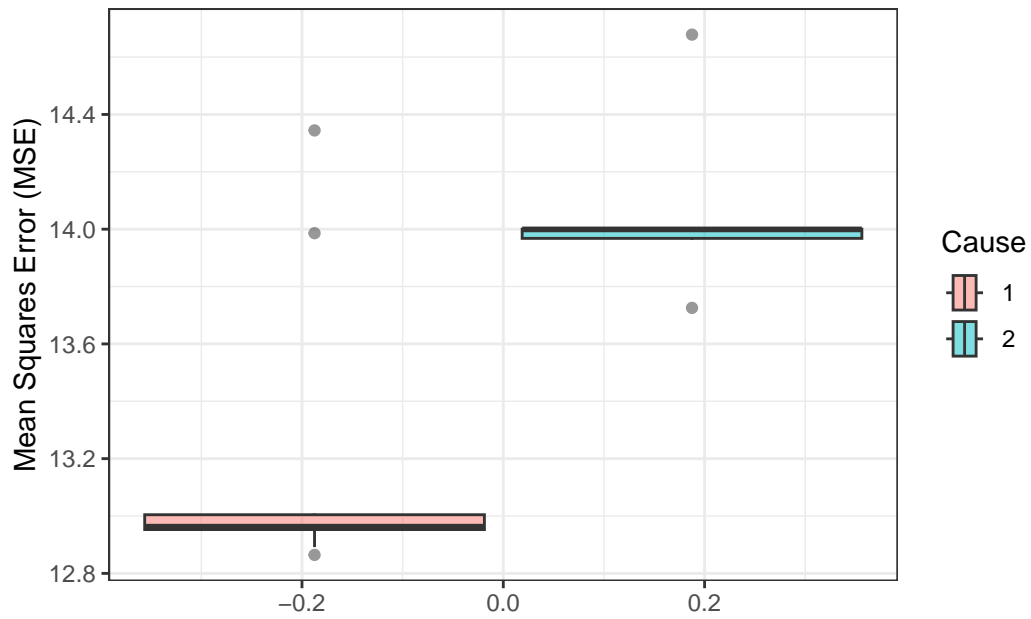
We consider a simple competing risks simulation setting for this example with the failure times for both causes being generated from an exponential distribution. The censoring times were generated from a uniform $Unif(1, 3)$ distribution which leads to an approximate censoring rate of $\sim 30\%$.

We consider a training/test/validation split of 0.60/0.20/0.20. The multinomial deviance was used as the criterion to select the optimal λ and the model was fit to the validation set using this value. The penalty value was selected from a grid of 500 values ranging from 0.001 to 0.5 spaced by 0.001. An elastic-net penalty ($\alpha = 0.5$) was set apriori for the model runs. The simulation was repeated 10 times to generate the results.

Simulation Results

For simulation results examine only the variable selection properties of the model, i.e the proportion of correctly selected variables, the proportion of incorrectly selected variables and the MSE of the estimated coefficients.





Next Steps

- Repeat simulation 100 times on compute canada with cross-validation to tune the penalty
- Repeat with LASSO penalty
- Fit CoxBoost and penalized Fine-Gray model for comparison